

Marcos Bosi Ferraz

TRADUÇÃO PARA O PORTUGUÊS E VALIDAÇÃO DO QUESTIONÁRIO
PARA AVALIAR A CAPACIDADE FUNCIONAL
"STANFORD HEALTH ASSESSMENT QUESTIONNAIRE"

Tese de Doutorado apresentada
à Escola Paulista de Medicina,
Curso de Pós-Graduação, Área
de Reumatologia

Orientador: Prof. Dr. Edgard Atra

São Paulo - Brasil

1990

À minha esposa Luci
pelo amor, amizade e incentivo.

À minha filha Beatriz
na expectativa de um mundo melhor.

Aos meus pais
que muito lutaram pela minha
formação.

À minha irmã Carla
pela alegria de sua companhia.

AGRADECIMENTOS:

Ao Prof. Dr. Edgard Atra pela orientação, dedicação, constante incentivo e amizade que muito contribuíram para minha formação profissional.

Ao Prof. Dr. José Goldenberg pelo estímulo, amizade e confiança sempre demonstrados.

À Profa. Dra. Emilia Inoue Sato pelo apoio, companheirismo e sugestões ao longo deste trabalho.

Ao Prof. Dr. Peter Tugwell pela confiança, incentivo e ensinamentos em Epidemiologia Clínica.

Ao Prof. Dr. Charlie H. Goldsmith pela preciosa orientação e auxílio na análise estatística.

Às fisioterapeutas Leda Magalhães Oliveira e Cristina Maria Giannecchini pela inestimável colaboração na realização deste trabalho. A Leda Magalhães Oliveira meu especial agradecimento.

À terapeuta ocupacional Pola Maria Polí Araújo pela valiosa participação no presente estudo.

Aos professores de inglês Terrence Edward Hill, Roberto Vicente Cruz Thenudo Lessa, João Paulo Almeida e Ana Cristina Soares pela preciosa colaboração na tradução do questionário.

Ao pós-graduando Dr. Rocyr André Visinoni pelo valioso auxílio na execução do ensaio clínico controlado randomizado.

À Professora Greta Alvarez Turcato pela providencial revisão ortográfica.

Aos pós-graduandos, residentes e estagiários da Disciplina de Reumatologia da Escola Paulista de Medicina pela amizade e constante apoio e incentivo.

A todos os pacientes pela compreensão e colaboração ao concordarem em participar deste estudo.

ÍNDICE:

I.	Introdução	1
II.	Objetivos	18
III.	Material e Métodos	20
IV.	Resultados	32
V.	Discussão	46
VI.	Conclusão	63
VII.	Summary	65
VIII.	Referências Bibliográficas	67

I. INTRODUÇÃO

O pequeno número de referências bibliográficas existentes na literatura médica da antiguidade de doenças algo semelhantes à artrite reumatóide (AR) sugere que a AR seja uma doença de origem relativamente recente (1). Além da literatura médica, outras fontes que poderiam auxiliar na determinação da primeira descrição ou menção de uma determinada doença são: o estudo da paleontologia humana, a pesquisa da literatura não científica e as obras de arte do passado.

Enquanto doenças reumáticas outras como osteoartrose, espondilite anquilosante, osteoporose e gota têm sido identificadas desde alguns séculos antes de Cristo, não existe entretanto nenhum dado convincente da existência de relatos de casos com AR (2). Entre as possíveis razões que poderiam explicar a ausência de evidências paleontológicas pode-se citar o exame pouco cuidadoso dos pequenos ossos das mãos e pés, a presença de outras doenças semelhantes à AR em alguns aspectos, como por exemplo, artrite psoriática e gota, bem como a precária preservação ao longo dos tempos de tecido sinovial proliferativo, característico da sinovite reumatóide (2).

O primeiro relato que destaca a AR como uma nova entidade foi apresentado por Augustin-Jacob Landrè-Beauvais em 1800, em sua tese de doutorado intitulada "Should we recognize a new type of Gout, to be called Primary Asthenic Gout?". O termo AR, no entanto, só foi proposto em 1859 por

Sir Alfred Baring Garrod, e a primeira caracterização clínica da doença somente foi publicada em 1890 por seu filho, Sir Archibald Edward Garrod (2). Descrições suplementares da AR foram posteriormente publicadas por Heberden, Haggert, Brodie, e outros (3).

Em 1952, Snorransen pela primeira vez sugere que a AR pudesse ser uma doença de origem recente, decorrente de algumas alterações ambientais que poderiam ter acontecido ao longo dos anos (4).

Atualmente, a AR é definida como uma doença inflamatória crônica, freqüentemente progressiva e de etiologia desconhecida. Caracteriza-se primariamente por um acometimento articular simétrico e bilateral, comprometendo freqüentemente as articulações das mãos e pés; embora as manifestações clínicas na maioria dos casos decorra do processo inflamatório articular com conseqüente perda funcional; as manifestações sistêmicas não são infreqüentes.

A prevalência da AR definida varia de 0.1% (5) a 3.0% (6). É uma doença encontrada em todo mundo, sendo que sua prevalência varia pouco de população para população quando critérios diagnósticos semelhantes são aplicados. A prevalência da AR aumenta discretamente em faixas etárias mais avançadas, sendo tal fato observável até os 70 anos. Até recentemente não havia nenhum dado publicado quanto à prevalência da AR no Brasil. Em 1988-1989, Emilia I. Sato e cols. (7), pela primeira vez, estudaram a prevalência da AR

numa comunidade japonesa da cidade de Mogi das Cruzes, São Paulo. A população estimada desta cidade é de 270.000 habitantes, sendo que aproximadamente 20.000 são de origem japonesa. Participaram do estudo 2.996 indivíduos, sendo a prevalência da AR definida em indivíduos com mais de 16 anos de 0.20% . Ainda, em estudo realizado em 3 capitais (Fortaleza, Rio de Janeiro e Porto Alegre), a AR foi a terceira causa de consultas médicas em serviços de reumatologia; a primeira foi osteoartrose (46.9%) e a segunda fibrosite (20.1%) (8).

Quanto ao sexo, acomete aproximadamente 3 mulheres para cada homem (9). Em adultos jovens há um maior acometimento do sexo feminino como tem sido observado em alguns estudos (10,11). A acrescentar, ainda não há uma explicação satisfatória pelo aparente maior acometimento do sexo feminino. A faixa etária de início da doença varia de 20 a 60 anos sendo o pico de início mais freqüentemente observado entre os 35 e 45 anos. Todas as raças parecem ser suscetíveis ao desenvolvimento da AR (12).

Quanto à etiopatogenia, apesar do grande número de pesquisas científicas, os motivos pelos quais um indivíduo desenvolve AR ainda permanecem desconhecidos. Um problema adicional no estudo da etiopatogenia da AR reside no fato de ser uma doença unicamente humana, e embora algumas formas experimentais existam, tais modelos não são amplamente satisfatórios.

Baseado em características clínicas e laboratorias, um conjunto de critérios para o diagnóstico da AR foi proposto e subseqüentemente revisado em 1987 por um comitê da Associação Americana de Reumatismo (ARA), atualmente chamado Colégio Americano de Reumatologia (ACR). A presença de 4 ou mais critérios proporciona uma sensibilidade de 93.5% e especificidade de 89.3% no diagnóstico da AR, quando aplicados em centros reumatológicos (13).

O tratamento de pacientes portadores de AR ainda constitui-se em um desafio, mesmo para os mais experientes reumatologistas. Como objetivos principais do tratamento podemos citar: o alívio dos sintomas, a interrupção do processo destrutivo tissular, a prevenção de deformidades articulares, bem como a restauração da capacidade funcional e a prevenção de acometimento de outros órgãos ou sistemas (14).

A presença, no entanto, de múltiplas opções terapêuticas, a ausência de um tratamento único reconhecidamente eficaz, a flutuação espontânea do curso da doença e a multiplicidade, bem como ausência de critérios clínicos amplamente aceitos para avaliar a atividade da doença contribuem sobremaneira para dificultar o estudo e avaliação de medidas terapêuticas eficazes.

Diversos parâmetros clínicos têm sido utilizados em ensaios clínicos envolvendo pacientes com AR, tais como: número de articulações inflamadas, número de articulações

dolorosas, índice articular de Ritchie (15), índice articular de Landsbury (16), duração da rigidez matinal, força de preensão, circunferência das articulações interfalangeanas proximais, tempo de caminhada de 15 metros, avaliação subjetiva da atividade clínica pelo paciente ou pelo médico, avaliação da dor, entre outros. Estes parâmetros tradicionais têm, no entanto, algumas limitações, uma vez que refletem em sua maioria aspectos isolados que constituem a sintomatologia do paciente. Assim, o significado clínico de uma melhora em cada um destes parâmetros é de tradução e compreensão difícil, ou seja, uma melhora de 30 minutos na rigidez matinal, uma redução de 3 articulações edemaciadas, um alívio de 20% da dor ou um aumento de 25mmHg na força de preensão podem não ter qualquer significado clínico na capacidade de execução das atividades diárias do paciente. Baseado neste fato, uma outra dimensão de parâmetro clínico tem sido proposta na avaliação de diversas modalidades terapêuticas na AR.

O valor da avaliação da capacidade ou habilidade funcional em pacientes com AR é reconhecido desde 1949 quando Steinbrocker e colaboradores propuseram a classificação funcional da Associação Americana de Reumatismo (17).

Apenas 16 a 20% dos pacientes com AR não apresentam evidências de diminuição da capacidade funcional (18). Por outro lado, 17% dos pacientes com AR são incapazes de manter

o cuidado e a higiene pessoal (18,19). Esta diminuição da capacidade funcional é habitualmente associada a uma maior dependência física e mental, bem como diminui a capacidade de trabalho e o rendimento financeiro, e aumenta a utilização dos serviços de assistência à saúde (20-23). Por estes aspectos é importante frisar que a prevenção ou o retardo na perda da capacidade funcional constitui-se em um dos principais objetivos da terapêutica anti-reumática.

A perda da capacidade ou habilidade funcional pode ser definida como qualquer restrição ou perda da capacidade de execução de atividades ou tarefas diárias consideradas normais ao cotidiano humano. É caracterizada pela deficiência no desempenho de uma atividade ou conduta, podendo ser temporário ou permanente, reversível ou irreversível e progressivo ou não (24).

Assim, em ensaios clínicos, com a finalidade de se avaliar de forma global a influência benéfica ou não de um determinado tratamento na vida diária dos pacientes, é importante que seja incluído entre os parâmetros clínicos principais um instrumento que seja capaz de avaliar globalmente o estado de saúde, a qualidade de vida e, ou especificamente a capacidade funcional dos pacientes; conseqüentemente este parâmetro clínico seria utilizado com um propósito de avaliação.

I.A. INSTRUMENTOS PARA AVALIAR A QUALIDADE DE VIDA

Um grande número de instrumentos, métodos e técnicas têm sido propostos para avaliar a qualidade de vida ou a capacidade funcional de pacientes com as mais diversas doenças. Podemos categorizar tais instrumentos em genéricos e específicos.

I.A.1. Instrumentos Genéricos:

Os instrumentos genéricos são aplicáveis a uma grande variedade de populações em virtude de incluírem aspectos relativos à função, disfunção e desconforto emocional e físico, sendo estes relevantes à qualidade de vida dos pacientes. Podem ainda ser subdivididos em 2 categorias: os que avaliam o estado de saúde e as medidas de "utility", termo este que não apresenta a conotação de utilidade de um estado de saúde, mas sim de preferência por um determinado estado de saúde.

I.A.1.a. Instrumentos para avaliar o Estado de Saúde:

Este tipo de instrumento proporciona uma avaliação de diferentes aspectos referentes à qualidade de vida do paciente. Regra geral possuem um sistema de contagem de pontos que são somados proporcionando assim uma nota final ou índice. Como o próprio nome o designa, são instrumentos criados para serem utilizados em uma grande variedade de patologias. Como exemplo, podemos citar um dos instrumentos

mais conhecidos para avaliar o perfil de saúde, ou seja, o "Sickness Impact Profile" (SIP) (25). Este instrumento tem sido utilizado em estudos envolvendo pacientes em reabilitação cardíaca (26), artroplastia total de articulação coxofemural (27), tratamento de lombalgia (28), entre outros. Os instrumentos "Nottingham Health Profile" (29) e McMaster Health Index Questionnaire" (30) são outros exemplos.

Instrumentos para avaliar o estado de saúde oferecem diversas vantagens, tais como, permitem a avaliação de modalidades terapêuticas em diversos aspectos referentes à qualidade de vida dos pacientes com o uso de um único instrumento, o que diminui o tempo de avaliação por parte do investigador. Outro aspecto interessante constitui-se no fato de permitirem a comparação dos efeitos de uma determinada intervenção terapêutica em diferentes patologias. Estes instrumentos, no entanto, apresentam algumas limitações, ou seja, por serem de avaliação global, não permitem o estudo de determinados aspectos particulares de uma patologia. Assim, o seu uso inadequado ocasionará uma não mensuração de aspectos importantes para uma determinada patologia, que na realidade são clinicamente importantes para o paciente. Por exemplo, no caso de uma intervenção terapêutica, o instrumento pode não ser capaz de detectar uma melhora (ou piora) clínica quando na realidade esta ocorreu e é importante na vida diária do paciente.

I.A.1.b. Índices para avaliar "utility":

Estes índices, originários e até recentemente só utilizados em estudos da área de economia, refletem a preferência dos pacientes por determinados estados de saúde. A qualidade de vida é medida e expressa através de um simples número que varia de 0.0 (morte) a 1.0 (saúde perfeita). Dois métodos podem ser utilizados na avaliação destes índices: 1. O investigador interroga o paciente com questões referentes a sua capacidade funcional e atividades relacionadas à vida diária e classifica o paciente em uma determinada categoria baseado em suas respostas. Cada uma destas categorias apresenta um valor previamente estabelecido (de 0 a 1) por outro grupo de pacientes e/ou indivíduos sadios. 2. Três ou mais cenários distintos que descrevem diferentes graus de atividade e gravidade da doença, e conseqüentemente diferentes limitações funcionais são apresentados aos pacientes. Tais cenários abordam ainda aspectos referentes ao comprometimento emocional, social e econômico decorrentes da doença. O investigador solicita ao paciente para que o mesmo de um valor entre 0 e 1 para cada cenário e a seguir para o seu próprio estado de saúde ou sua qualidade de vida. Três diferentes técnicas podem ser utilizadas neste método, ou seja: escala visual, escolha pelo tempo ("Time Trade-off") ou escolha pela chance ("Standard Gamble") (31,32).

Uma vantagem na utilização de índices para avaliar

"utility" constitui-se na possibilidade de proporcionar a cada paciente uma avaliação de estados de saúde ou situações que são particularmente importantes para ele, podendo tais situações serem pontuadas adequadamente. Assim, se para um determinado paciente a locomoção é extremamente importante e este é paraplégico, o mesmo pode determinar em números o que esta paraplegia esta acarretando em sua qualidade de vida. Nos índices anteriormente citados tal paraplegia corresponderia, por exemplo, quando expresso em números a apenas 10% do valor final de seu estado de saúde ou qualidade de vida. Os outros 90% avaliariam aspectos outros tais como função emocional, intelectual, dor, recreação, etc. Uma desvantagem presente em tais índices constitui-se em sua relativa complexidade, bem como questionável validade da caracterização do estado de saúde ou qualidade de vida em uma escala simplificada de 0.0 a 1.0, não apresentando desta forma um significado clínico próprio.

I.A.2. Instrumentos Específicos:

Um modo alternativo de avaliar a qualidade de vida dos pacientes constitui-se no uso de instrumentos capazes de avaliar individualmente e especificamente determinados aspectos da qualidade de vida. Tal procedimento proporciona uma maior capacidade de detecção de melhora ou piora deste aspecto específico em estudo; ou seja, o instrumento apresenta uma maior suscetibilidade à alteração, propriedade

esta muito importante na avaliação de modalidades terapêuticas. Estes instrumentos podem ser específicos para uma determinada doença (por exemplo: AR, coronariopatia, diabetes mellitus), serem aplicáveis a uma determinada população de pacientes (pacientes idosos), ou então, específicos a uma determinada função (função sexual). Instrumentos para serem utilizados em doenças específicas têm se mostrado de grande valor na avaliação de modalidades terapêuticas em ensaios clínicos (33-35). Uma desvantagem peculiar a este grupo de instrumentos reside no fato de não proporcionarem uma comparação entre as diferentes patologias.

Os índices, instrumentos ou questionários para avaliar a qualidade de vida podem ainda ser administrados por entrevistadores (médico(a)s, biomédico(a)s ou enfermeiro(a)s) ou serem auto-administráveis. Ambos os modos apresentam vantagens e desvantagens, a citar: entre as desvantagens dos instrumentos auto-administráveis pode-se citar a exclusão de alguns grupos de pacientes (analfabetos, por exemplo), a maior chance de obtenção de dados incompletos, e a possibilidade de ocasionar certa ansiedade nos pacientes. Entre as desvantagens da administração do instrumento ou questionário por entrevistador podem ser citadas a necessidade de um treinamento ou padronização prévia da entrevista, o fato de ser de certa forma caro, além do que as características do entrevistador (idade,

sexo, raça, etc) podem influenciar a obtenção dos resultados.

I.A.3. Avaliação da qualidade de vida em reumatologia:

Em reumatologia, diversos instrumentos ou índices têm sido propostos e utilizados com o intuito de avaliar a qualidade de vida dos pacientes, sendo alguns destes genéricos, outros específicos para uma determinada patologia ou para uma determinada função. Entre estes, podemos citar: "Taylor's index" (36), "British Empire Rheumatism Council Scheme" (37), "Donaldson Activities of daily living (ADL) instrument" (38), "Katz ADL instrument" (39), "Lee's Functional status instrument" (40), "Keitel index" (41), "Barthel index" (42), "Convery scale" (43), "Functional Status Index (FSI)" (44), "Stanford Health Assessment Questionnaire (HAQ)" (45), "Index of Well-Being" (46), "Arthritis Measurement Impact Scale (AIMS)" (47), "McMaster Health Index Questionnaire" (30), "Mactar Patient Preference Questionnaire" (48) and "EPM-ROM scale" (49).

Como pode ser observado, há um grande número de instrumentos capazes de avaliar a qualidade de vida ou perda da capacidade física, mental, intelectual ou emocional de pacientes com AR. A seleção, no entanto, de um determinado instrumento para ser utilizado em uma pesquisa clínica deve ser orientada pelos objetivos desta pesquisa, bem como pela proposição de uso deste instrumento, ou seja, com o intuito

de avaliação, de predição ou discriminação. Assim, um instrumento ao ser construído para a finalidade de discriminação, pode não possuir as propriedades de medida necessárias para a predição ou avaliação da eficácia de modalidades terapêuticas e vice-versa. Outro aspecto importante a ser considerado constitui-se na caracterização da população a ser estudada; neste caso, por exemplo, um instrumento auto-administrável não pode ser utilizado em pesquisas envolvendo uma população de pacientes analfabetos ou com condições intelectuais mínimas para sua compreensão. Ainda devem ser consideradas e avaliadas as propriedades de medida do instrumento selecionado (reprodutibilidade, validade, suscetibilidade à alteração), bem como sua praticidade na administração, tanto em termos de tempo, quanto ao custo da mesma; seu escore final deve ser de fácil obtenção, compreensível e confiável. Instrumentos complexos requerem entrevistadores diferenciados e treinados na sua administração, o que sem dúvida os tornam economicamente inviáveis na maioria das oportunidades.

É importante que seja ressaltado que o instrumento a ser selecionado respeite os preceitos éticos da pesquisa.

I.B. ASPECTOS ENVOLVIDOS NA TRADUÇÃO DE INSTRUMENTOS

Todo projeto de pesquisa transcultural envolve o uso da linguagem e conseqüentemente, problemas de comunicação podem ocorrer com a tradução de uma simples sentença (50).

Problemas de equivalência em traduções são freqüentemente observados. Secherest e cols. (50) descrevem 4 tipos de equivalência que, de certa forma, nas traduções apresentam diferentes tipos de efeitos e conseqüências, ou seja: equivalência de vocabulário, equivalência idiomática, equivalência de gramática e sintaxe, e equivalência experimental. Estes problemas são mais freqüentemente observáveis tanto quanto mais distintas forem as características culturais e linguísticas.

I.B.1. Equivalência de Vocabulário:

A tradução de qualquer palavra pode teoricamente ser solucionada com o auxílio de um dicionário, no entanto, em dicionários, a maioria das palavras são definidas por vários termos ou palavras sinônimas, não sendo fácil portanto a identificação do termo ou palavra mais adequada para a tradução. A acrescentar ainda a presença de diferenças culturais e sociais, bem como tradições linguísticas que tornam a tradução de determinadas palavras impróprias ou impossíveis (50). Como exemplo, podem ser citadas as palavras em inglês "outcome" e "cohort" ainda sem uma tradução adequada para o português.

I.B.2. Equivalência idiomática:

Dificuldades na tradução de expressões idiomáticas são freqüentemente observadas, sendo a equivalência perfeita

destas expressões raramente alcançadas. Como exemplo, podemos citar a expressão coloquial em inglês "I am blue" que significa estar deprimido. Sua tradução para o português, "eu estou azul", no entanto, não apresenta um significado maior ou pode ser entendida popularmente como se tudo estivesse bem com esta pessoa.

I.B.3. Equivalência de gramática e sintaxe:

Tanto a gramática quanto a sintaxe se constituem em aspectos críticos na correta compreensão de qualquer tradução. Como exemplo, podemos citar as formas verbais, ou seja, o tempo verbal "present perfect", um problema de equivalência gramatical observado em diferentes línguas (50).

I.B.4. Equivalência experimental ou cultural:

Para que uma tradução transcultural alcance seu objetivo é importante que os termos traduzidos e utilizados façam alusão à experiências que sejam familiares a ambas culturas. Werner e Campbell (51) chamam este aspecto de "tradução cultural" para distingui-la da tradução lingüística. Assim, se duas culturas apresentam grandes diferenças quanto aos seus arranjos sociais, seus modos de vida e valores, é extremamente difícil obter-se a equivalência cultural, mesmo com uma tradução cuidadosa do ponto de vista lingüístico (50).

Como a maioria dos instrumentos e questionários para avaliar a qualidade de vida foram propostos e publicados por autores de outras línguas que não o português, os aspectos acima abordados devem ser considerados na tradução destes instrumentos para a língua portuguesa.

II. OBJETIVO:

O objetivo do presente estudo é identificar um instrumento ou questionário para avaliar a capacidade ou habilidade funcional de pacientes com artrite reumatóide, traduzi-lo para o português e finalmente avaliar suas propriedades de medida após sua tradução, validando desta forma a versão em português do instrumento escolhido. O instrumento a ser selecionado será utilizado na avaliação de modalidades terapêuticas envolvendo pacientes com AR.

III. MATERIAL E MÉTODOS

III.A. IDENTIFICAÇÃO E SELEÇÃO DO INSTRUMENTO

Em virtude da quase totalidade dos instrumentos existentes serem publicados na língua inglesa (originariamente ou traduzidos), a literatura médica de língua inglesa publicada nos últimos 30 anos será revisada com a finalidade de identificar instrumentos ou questionários para avaliar a capacidade ou habilidade funcional. Nesta revisão as seguintes fontes serão utilizadas: Index Medicus, sistema MEDLINE e Citation Index, bem como a referência dos artigos localizados será cuidadosamente observada visando a identificação de outros instrumentos não localizados pelos métodos anteriores.

O instrumento ou questionário a ser traduzido para o português será utilizado com o propósito de avaliar diferentes modalidades terapêuticas na AR. Assim, este instrumento além de avaliar a capacidade ou habilidade funcional de pacientes com AR, já deve ter sido utilizado e aceito pela comunidade científica internacional e já ter tido suas propriedades de medida avaliadas. Por ser um instrumento avaliador as seguintes propriedades de medida deverão ser consideradas na sua seleção: reprodutibilidade, validade construtiva longitudinal (relação entre a alteração constatada pelo instrumento e as alterações observadas pelos outros parâmetros de medida da atividade clínica, decorrentes da ação de uma intervenção terapêutica) e suscetibilidade à alteração ("responsiveness") (capacidade

do instrumento detectar uma melhora ou piora clínica decorrente de uma intervenção terapêutica, quando esta realmente existe) (52). Deve ainda ser de fácil administração e de obtenção de resultado final não complicado e compreensível.

III.B. STANFORD HEALTH ASSESSMENT QUESTIONNAIRE (HAQ)

O HAQ (45) foi um instrumento inicialmente elaborado com a finalidade de permitir a avaliação de parâmetros de medida em ensaios terapêuticos envolvendo pacientes com AR e outras doenças reumáticas. Fries e cols. (45) relatam que as 5 dimensões importantes para um paciente reumático são o desejo de uma longa sobrevida, viver sem dor, ter capacidade para execução das tarefas da vida diária sem limitação, não apresentar toxicidade com os tratamentos propostos e ser capaz de trabalhar e ser independente do ponto de vista econômico-financeiro (45). Estas 5 dimensões compõem o HAQ, no entanto, os escores obtidos em cada uma destas dimensões, por abordarem aspectos distintos, não podem ser somados. Assim, apesar do HAQ ser considerado um instrumento único, este constitui-se em 5 instrumentos independentes que quando utilizados em conjunto permitem uma avaliação dos mais importantes aspectos inerentes a doença reumática.

Um fato a ser realçado, no entanto, consiste na observação de que desde a publicação do trabalho de proposição inicial do instrumento até a presente data um

grande valor e atenção tem sido dado à dimensão capaz de avaliar a capacidade e habilidade funcional do HAQ. As outras dimensões, tais como, a avaliação da dor ou do aspecto econômico, embora apresentando uma estrutura adequada, tem sido esporadicamente citados e utilizados em trabalhos científicos já realizados. A salientar ainda que a dimensão capacidade e habilidade funcional foi a primeira a ser criteriosamente desenvolvida e avaliada quanto as suas propriedades de medida, fato este de extrema importância para sua divulgação no meio científico. Assim, em decorrência do fato acima exposto, a dimensão capacidade e habilidade funcional tem sido reconhecida como o próprio HAQ.

O HAQ é um instrumento que pode ser auto-administrável ou administrado por entrevistadores composto por 8 componentes que avaliam aspectos distintos da atividade da vida diária. Cada um destes componentes por sua vez apresenta 2 ou 3 questões relacionadas às atividades físicas. A dificuldade relatada pelo paciente em executar uma determinada atividade física na semana que antecede a entrevista é dada um nota que varia de 0 (sem qualquer dificuldade) a 3 (atividade incapaz de ser executada pelo paciente). A nota do componente equivale a maior nota observada em uma das 2 ou 3 questões. O escore final do instrumento é obtido através da média aritmética das notas

dos 8 componentes; assim, o escore final do instrumento varia de 0 a 3. Porém, quando em alguma das questões dos componentes as atividades físicas são executadas com o auxílio de apoio, bengala ou órteses, a nota do referido componente é elevada automaticamente para 3.

O estudo das propriedades de medida do HAQ, ou seja, sua reprodutibilidade, validade e suscetibilidade à alteração ("responsiveness") estão amplamente demonstradas, o que justifica seu amplo uso na área da pesquisa reumatológica, tornando-o um dos instrumentos mais conhecidos e valorizados na pesquisa médica. Vale a pena ainda ser ressaltado que tal instrumento é de fácil administração sendo que a entrevista não leva mais do que 10 a 15 minutos para ser concluída.

Pelos motivos expostos acima, o presente instrumento será selecionado para ser traduzido para o português e validado após a sua tradução.

III.C. TRADUÇÃO PARA O PORTUGUÊS

Em virtude do presente estudo ser um exemplo de pesquisa transcultural, a tradução do instrumento ou questionário selecionado deverá obedecer os seguintes estágios: 1. Tradução inicial; 2. Avaliação da tradução inicial; e 3. Avaliação da equivalência cultural.

III.C.1. Tradução inicial:

O HAQ será inicialmente traduzido para o português por 2 professores de inglês cientes dos objetivos desta pesquisa. No caso de divergências nas traduções os 2 professores as discutirão devendo os mesmos chegar a um consenso quanto à tradução inicial.

III.C.2. Avaliação da tradução inicial:

O método escolhido para avaliar a tradução inicial para o português será a versão desta tradução para o inglês e comparação com o instrumento original. Dois outros professores de inglês, não conhecedores do instrumento em inglês, participarão deste estágio. As discrepâncias existentes serão documentadas e analisadas por um grupo composto por 2 reumatologistas, 2 fisioterapeutas, 1 terapeuta ocupacional e por 2 professores de inglês (sendo 1 de cada estágio). Quando necessário as sentenças serão re-escritas até a obtenção de um consenso.

III.3.C. Avaliação da equivalência cultural:

Em se tratando de um instrumento para avaliar a capacidade ou habilidade funcional é importante que o instrumento ou componentes deste instrumento apresentem ou avaliem experiências reais vivenciadas por ambas culturas. Assim, uma alternativa de resposta será acrescentada ao

instrumento já vertido para o português, ou seja, a cada uma das 20 perguntas presentes no HAQ será dada a opção de resposta "não aplicável". Com a finalidade de identificar quais as atividades físicas que seriam culturalmente não compatíveis ou não executadas com certa regularidade pelos pacientes brasileiros com doenças reumáticas, um grupo de pacientes com diagnóstico de AR, pelos critérios da ARA de 1987, será selecionado randomicamente e entrevistado. Os pacientes serão então orientados quanto aos objetivos da pesquisa e sobre a categoria de resposta "não aplicável". Questões que tiverem mais que 20% de respostas "não aplicável" serão consideradas como impróprias e passíveis de substituição. Após a identificação das mesmas um segundo grupo composto por 5 pacientes com AR, 2 reumatologistas, 2 fisioterapeutas e 1 terapeuta ocupacional se reunirá para selecionar as questões substitutivas, sem no entanto alterar de modo significativo a estrutura e propriedades físicas das questões tidas como impróprias.

Ainda a acrescentar, com as modificações passíveis de serem realizadas, é importante que a versão em português do HAQ não apresente questões redundantes ou repetitivas. Para tanto, cada componente individual da versão em português do HAQ será correlacionado com os outros componentes e na eventualidade de observação de alta ou muito baixa correlação o(s) componente(s) será(ão) eliminado(s).

III.D. AVALIAÇÃO DAS PROPRIEDADES DE MEDIDA DA VERSÃO EM PORTUGUÊS DO HAQ

III.D.1. Reprodutibilidade:

A reprodutibilidade da versão em português do HAQ será avaliada através de 3 entrevistas. Pacientes com diagnóstico de AR, de acordo com os critérios da ARA de 1987, selecionados aleatoriamente do ambulatório de reumatologia da Escola Paulista de Medicina serão avaliados por 2 entrevistadores, a saber: entrevistador 1 no primeiro dia (1a. X) e no quinto dia (2a. X); no quinto dia uma avaliação adicional será feita, desta vez pelo entrevistador 2. Todas as entrevistas serão realizadas no período da manhã.

III.D.2. Validade:

Por se tratar de um instrumento a ser utilizado com a finalidade de avaliar a efetividade de uma determinada modalidade terapêutica, a validade do HAQ em português pode ser demonstrada através da constatação de uma correlação entre uma melhora ou piora dos valores obtidos pelo HAQ e outros parâmetros clínicos comumente utilizados em ensaios envolvendo pacientes com AR (validade construtiva longitudinal ou "longitudinal construct validity"). Assim, é necessário que seja demonstrado, por exemplo, que pacientes com uma redução de 50% do número de articulações inflamadas

ou com dor apresentem também uma redução na dificuldade para a execução das tarefas da vida diária. Em um paciente com AR é importante, portanto, demonstrar que as alterações observadas em determinado parâmetro clínico ao longo de um determinado período se relacione positivamente com as alterações observadas nos valores do HAQ em português. Assim, o modo mais convincente de se demonstrar a validade do HAQ consiste na avaliação prospectiva de um grupo de pacientes com AR passíveis de melhora ou piora clínica. Com a finalidade de se demonstrar a validade da versão em português do HAQ, o HAQ e outros parâmetros clínicos e laboratoriais serão sistematicamente avaliados em um ensaio clínico controlado randomizado para avaliar e comparar a efetividade de 2 doses distintas de metilprednisolona administradas sob a forma de pulsoterapia a pacientes com AR em atividade clínica. A escolha deste ensaio clínico como meio de validação do HAQ baseia-se no fato de que indivíduos com AR em fase de atividade clínica da doença, certamente apresentarão algum grau de melhora com tal intervenção, o que permitirá desta forma a correlação da melhora constatada pelo HAQ com a melhora observada através de outros parâmetros clínicos.

III.D.3. Suscetibilidade à alteração:

Suscetibilidade à alteração ("responsiveness") relaciona-se a capacidade que um parâmetro clínico possui

para detectar uma melhora ou piora clínica quando esta realmente ocorre. A suscetibilidade à alteração da versão em português do HAQ será obtida através da comparação da capacidade de detecção da melhora clínica constatada pela versão em português do HAQ com a capacidade de detecção desta melhora observada pelos demais parâmetros clínicos e laboratoriais utilizados no ensaio terapêutico supra-citado. Tal procedimento possibilitará, portanto, a determinação da suscetibilidade à alteração da versão em português do HAQ em relação a suscetibilidade à alteração dos demais parâmetros freqüentemente utilizados nos ensaios clínicos envolvendo pacientes com AR.

III.E. ENSAIO CLÍNICO CONTROLADO RANDOMIZADO

No ensaio clínico controlado randomizado duplo-cego para avaliar a efetividade de 2 doses de Metilprednisolona administrada na forma de pulsoterapia, pacientes com AR em atividade serão avaliados clínica e laboratorialmente nos tempos 0, 1, 3 e 6 semanas após a administração da droga. As doses a serem comparadas serão de 5mg/kg e 10mg/Kg de Metilprednisolona, sendo que a dose total deverá ser administrada endovenosamente diluída em 250 ml de soro fisiológico a 0.9% . Diversos parâmetros clínicos serão utilizados na avaliação clínica e laboratorial destes pacientes, a citar: força de preensão, rigidez matinal, dor, número de articulações inflamadas, escala EPM-ROM (49),

Questionário Mactar (48), a versão em português do HAQ, VHS, hemograma, entre outros. Em virude de ser um estudo duplo-cego e a comparação entre as 2 dosagens de metilprednisolona não ser de interesse a este estudo, a identificação dos pacientes no que se refere ao grupo a que forem randomizados será mantida em sigilo para não prejudicar o resultado final deste estudo. É importante desta forma frisar que a não identificação dos grupos de tratamento não interferirá em nada na avaliação da validade e suscetibilidade à alteração da versão em português do HAQ.

As razões, objetivos e critérios de elegibilidade de pacientes deste ensaio clínico controlado randomizado, bem como seus resultados e conclusões finais serão motivo de uma publicação posterior.

III.F. ANÁLISE ESTATÍSTICA

Análise estatística descritiva será utilizada na caracterização dos pacientes utilizados nas diferentes etapas deste estudo.

O coeficiente de correlação de Pearson será utilizado na avaliação da correlação entre os 8 componentes da versão em português do HAQ. Para a demonstração de ausência de redundância entre as diferentes questões e aceitação de que as mesmas avaliam de certa forma o mesmo parâmetro, ou seja, capacidade funcional, será necessário que as correlações entre os diversos componentes da versão em português do HAQ

variem de 0.200 a 0.800.

A reprodutibilidade intra-observador e inter-observador da versão em português do HAQ será avaliada através do coeficiente de correlação de Pearson. Coeficientes maiores que 0.8 serão considerados clinicamente importantes.

Na validação da versão em português do HAQ os valores dos diversos parâmetros clínicos e laboratoriais obtidos através da diferença pós-pulsoterapia (na 1a., 3a. e 6a. semanas) e pré-pulsoterapia serão correlacionados com os valores da diferença pós (na 1a., 3a. e 6a. semanas) e pré-pulsoterapia da versão em português do HAQ, utilizando-se mais uma vez o coeficiente de correlação de Pearson.

A suscetibilidade à alteração de um instrumento é diretamente proporcional à média das variações do resultado deste instrumento e inversamente proporcional a variabilidade das variações dos resultados entre os pacientes envolvidos no estudo (neste caso, o ensaio clínico comparando 2 doses da metilprednisolona); assim, a suscetibilidade à alteração será obtida através da razão: média da variação do instrumento/erro padrão da variação do instrumento. Esta razão será calculada para cada parâmetro clínico e laboratorial sendo as mesmas comparadas aos valores obtidos pela versão em português do HAQ. O nível de significância estatística será definido em 0.05.

IV. RESULTADOS

IV.A. TRADUÇÃO PARA O PORTUGUÊS

A versão em português do HAQ e sua tradução para o inglês está apresentada na Tabela 1. Trinta e um pacientes com AR participaram na identificação das questões sem equivalência cultural. O resultado da avaliação da equivalência cultural e a identificação das questões consideradas inapropriadas por mais de 20% da população estudada estão apresentados na Tabela 2. As questões de números 7, 11, 15, 19 e 20 foram substituídas por novas questões para adequar o questionário às condições sócio-culturais e econômica de nossa população. A versão em português do HAQ (Tabela 1) foi então administrada a outros 22 pacientes com AR para verificar tal adequação, sendo que as 20 questões foram consideradas apropriadas por mais de 90% dos pacientes.

IV.B. AVALIAÇÃO DAS PROPRIEDADES DE MEDIDA

IV.B.1. Reprodutibilidade:

A Tabela 3 apresenta as características clínicas e demográficas dos 38 pacientes com AR incluídos na avaliação da reprodutibilidade da versão em português do HAQ. A reprodutibilidade intra-observador e entre-observadores do escore global do instrumento, bem como de cada componente individual, avaliados pelo coeficiente de correlação de Pearson, está apresentada na Tabela 4. Os coeficientes de

Tabela 1: Versão em português do HAQ e sua versão para o inglês.

Você é capaz de: (Are you able to:)

C1—Vestir-se e cuidar-se (Dressing and grooming)

1. Vestir-se, inclusive amarrar os cordões dos sapatos e abotoar suas roupas?
(Dress yourself, including tying shoelaces and doing buttons?)
2. Lavar sua cabeça e seus cabelos?
(Shampoo your hair?)

C2—Levantar-se (Arising)

3. Levantar-se de maneira ereta de uma cadeira de encosto reto e sem braços?
(Stand up straight from an armless straight chair?)
4. Deitar-se e levantar-se da cama?
(Get in and out of bed?)

C3—Comer (Eating)

5. Cortar um pedaço de carne?
(Cut your meat?)
6. Levar a boca um copo ou uma xícara cheia de café, leite ou água?
(Lift a full cup or glass to your mouth?)
7. Abrir um saco de leite comum?
(Open an ordinary milk plastic bag?)

C4—Caminhar (Walking)

8. Caminhar em lugares planos?
(Walk outdoors on flat ground?)
9. Subir 5 degraus?
(Climb up 5 steps?)

C5—Higiene (Hygiene)

10. Lavar seu corpo inteiro e secá-lo após o banho?
(Wash and dry your entire body?)
11. Tomar um banho de chuveiro?
(Take a shower bath?)
12. Sentar-se e levantar-se de um vaso sanitário?
(Get on and off the toilet?)

C6—Alcançar (Reach)

13. Levantar os braços e pegar um objeto de aproximadamente 2.5 quilos que esta posicionado pouco acima da cabeça?
(Reach and get down a 5-pound object from just above your head?)
14. Curvar-se para pegar suas roupas no chão?
(Bend down to pick up clothing from the floor?)

C7—Preensão (Grip)

15. Segurar-se em pé no ônibus ou metrô?
(Stand up by yourself on a bus or a subway?)
16. Abrir potes ou vidros de conservas que tenham sido previamente abertos?
(Open jars which have been previously opened?)
17. Abrir e fechar torneiras?
(Turn faucets on and off?)

C8—Outras atividades (Other activities)

18. Fazer compras nas redondezas aonde mora?
(Run errands and shop?)
 19. Entrar e sair de um ônibus?
(Get onto and off a bus?)
 20. Realizar tarefas tais como usar a vassoura para varrer e rôdo para puxar a água?
(Do everyday tasks such as sweeping or mopping the floor?)
-

Tabela 2: Questões presentes no HAQ em inglês que foram consideradas "não aplicáveis" pelos pacientes brasileiros com artrite reumatóide.

Pacientes com Artrite Reumatóide (n=31)	
Questões do HAQ	número de pacientes que responderam "Não Aplicável" (%)
Q1	0 (0.0%)
Q2	1 (3.2%)
Q3	0 (0.0%)
Q4	0 (0.0%)
Q5	0 (0.0%)
Q6	0 (0.0%)
Q7	30 (96.7%)*
Q8	0 (0.0%)
Q9	0 (0.0%)
Q10	0 (0.0%)
Q11	30 (96.7%)*
Q12	1 (3.2%)
Q13	2 (6.4%)
Q14	0 (0.0%)
Q15	13 (43.2%)*
Q16	2 (6.4%)
Q17	0 (0.0%)
Q18	1 (3.2%)
Q19	12 (38.7%)*
Q20	15 (48.4%)*

* Questões identificadas que foram substituídas na versão em português do HAQ.

Tabela 3: Características demográficas e clínicas dos 38 pacientes com artrite reumatóide incluídos no estudo para avaliar a reprodutibilidade do HAQ.

Pacientes com artrite reumatóide (n=38)		
Sexo		
Feminino:masculino		33:5
Idade (anos)		
média (DP)		49.6 (13.8)
Duração da doença		
média (DP)		7.8 (7.4)
Rigidez matinal (min)		
média (DP)		79.2 (70.5)
Classe Funcional* casos (%)	I	6 (15.8)
	II	23 (60.5)
	III	7 (18.4)
	IV	2 (5.3)
Resultado do HAQ		
média (DP)		1.143 (0.689)

* De acordo com o critério de Steinbrocker (17).

Tabela 4: Reprodutibilidade intra-observador e inter-observador da versão em português do HAQ e de cada componente individual avaliados através do coeficiente de correlação de Pearson.

Pacientes com artrite reumatóide (n=38)			
Coeficientes de correlação (r)*			
HAQ	Média (DP)	Intra-observador	Inter-observador
Total	1.143 (0.689)	0.905	0.830
C ₁	0.846 (0.812)	0.785	0.725
C ₂	0.897 (0.680)	0.501	0.599
C ₃	1.103 (1.021)	0.850	0.822
C ₄	1.000 (0.795)	0.793	0.618
C ₅	0.974 (0.743)	0.613	0.604
C ₆	1.282 (0.857)	0.779	0.663
C ₇	1.436 (0.882)	0.760	0.749
C ₈	1.359 (0.873)	0.782	0.779

* todos os valores com $P < 0.01$

correlação de Pearson entre os diversos componentes da versão em português do HAQ, administrada pelo entrevistador 1 no tempo 1, estão apresentados na Tabela 5.

IV.B.2. Ensaio clínico controlado randomizado:

Vinte pacientes já foram randomizados, sendo que 19 destes pacientes completaram as 6 semanas de seguimento clínico. Um paciente não retornou aos seguimentos clínicos previstos, tendo somente comparecido à administração da metilprednisona. Nenhum efeito colateral foi observado ou relatado pelos pacientes. As características clínicas e demográficas destes pacientes estão apresentadas na Tabela 6.

IV.B.3. Validade:

Os coeficientes de correlação de Pearson entre as variações dos escores obtidos com a administração da versão em português do HAQ e as variações dos outros parâmetros clínicos dos 19 pacientes com AR avaliados entre as semanas 0 e 1, 0 e 3, e 0 e 6 semanas, decorrentes do uso da pulsoterapia com metilprednisona estão apresentados na Tabela 7. Podemos observar que as correlações entre as variações dos escores da versão em português do HAQ e as variações dos outros parâmetros clínicos foram clinicamente importantes e estatisticamente significantes, exceção feita a correlação entre a versão em português do HAQ e a VHS

Tabela 5: Coeficientes de correlação de Pearson entre os componentes da versão em português do HAQ, quando avaliados pelo observador 1 no tempo 1.

Coeficientes de correlação (r) (n=38)								
	C ₁	C ₂	C ₃	C ₄	C ₅	C ₆	C ₇	C ₈
C ₁	1	.590	.559	.774	.647	.631	.610	.599
C ₂		1	.432	.633	.567	.412	.383	.285
C ₃			1	.389	.594	.538	.592	.312
C ₄				1	.535	.502	.402	.455
C ₅					1	.590	.620	.542
C ₆						1	.529	.459
C ₇							1	.714
C ₈								1

Tabela 6: Características clínicas e demográficas dos 19 pacientes incluídos no ensaio clínico controlado para avaliar e comparar a efetividade de duas doses de metilprednisolona administradas na forma de pulsoterapia.

Pacientes com artrite reumatóide (n=19)		
Sexo		
Feminino:masculino		18:1
Idade (anos)		
média (DP)		44.4 (11.6)
Duração da doença		
média (DP)		5.6 (4.4)
Rigidez matinal (min)		
média (DP)		151.8 (99.8)
Dor (escala numérica)		
média (DP)		6.0 (2.0)
Força de preensão (mmHg)		
média (DP)		59.5 (23.7)
Número de articulações inflamadas, média (DP)		
		17.9 (8.2)
Classe Funcional*		
II	14	(73.6)
casos (%)		
III	5	(26.4)

* De acordo com o critério de Steinbrocker (17).

Tabela 7: Coeficientes de correlação de Pearson entre a variação do resultado global da versão em português do HAQ e a variação dos diversos parâmetros clínicos observados entre as semanas 0 e 1, 0 e 3, e 0 e 6 em 19 pacientes com AR em atividade tratados com metilprednisolona na forma de pulsoterapia.

Parâmetros	Coeficiente de correlação de Pearson	Prob.
HAQ X Dor	0.446	P<0.001
HAQ X EPM-ROM	0.634	P<0.001
HAQ X Rigidez matinal	0.667	P<0.001
HAQ X Núm. artic. inflam.	0.483	P<0.001
HAQ X força de preensão	0.585	P<0.001
HAQ X Mactar	0.540	P<0.001
HAQ X VHS	-0.113	NS

(NS). Estes dados constataam a validade construtiva longitudinal ("longitudinal construct validity") da versão em português do HAQ, ou seja, a melhora clínica observada pela versão em português do HAQ correlaciona-se positivamente com a melhora clínica observada por outros parâmetros clínicos comumente utilizados em ensaios clínicos envolvendo pacientes com AR. Dois aspectos adicionais devem ainda ser considerados ao se avaliar a validade de um instrumento, ou seja: a validade aparente ("face validity"), isto é, a verificação se o instrumento avalia a dimensão teoricamente proposta e de interesse, que no caso do HAQ se caracteriza pela avaliação da capacidade ou habilidade funcional; e a validade de conteúdo ("content validity") que verifica se o instrumento como um todo permite uma ampla avaliação da dimensão proposta. Ambos aspectos, embora de demonstração subjetiva, podem ser satisfatoriamente observados na versão em português do HAQ.

IV.B.4. Suscetibilidade à alteração:

Com a finalidade de avaliar a suscetibilidade à alteração da versão em português do HAQ, a variação dos parâmetros clínicos entre os tempos 0 e 1 semana foram utilizados. A melhora média (e desvio padrão) observada nos diversos parâmetros clínicos e laboratoriais na 1a. semana após a administração endovenosa de metilprednisolona 5mg/kg ou 10mg/kg na forma de pulsoterapia está apresentada na

Tabela 8. Ao se utilizar a fórmula proposta é possível estabelecer uma comparação entre os diversos parâmetros clínicos. Assim, a suscetibilidade à alteração de cada parâmetro está apresentada na Tabela 9; quanto maior o valor obtido, maior a suscetibilidade à alteração do parâmetro, ou seja, maior a sua capacidade de detectar uma melhora clínica quando esta realmente existe. Podemos desta forma observar que a versão em português do HAQ apresentou um valor bem superior aos parâmetros dor, força de preensão, Questionário Mactar (48) e VHS. Conforme o esperado, a maior suscetibilidade à alteração foi observada com o parâmetro número de articulações inflamadas, parâmetro este reconhecidamente tido como o principal em diversos ensaios clínicos envolvendo pacientes com AR.

Finalizando, os dados acima apresentados demonstram que a versão em português do HAQ constitui-se em um instrumento responsivo.

Tabela 8: Melhora clínica média (tempo 1 - tempo 0) observada, em 19 pacientes, através dos diversos parâmetros ao término da primeira semana após a administração endovenosa de metilprednisolona.

Parâmetro Clínico	Média (desvio padrão) (n = 19)
Dor (escala numérica)	- 0.944 (1.617)
EPM-ROM	- 1.700 (2.902)
Rigidez matinal (min)	- 76.3 (68.3)
Força de preensão (mmHg)	10.7 (19.8)
Número de artic. inflam.	- 4.7 (4.3)
Mactar	- 32.6 (32.2)
VHS (mmHg)	- 19.3 (20.5)
HAQ em português	- 0.570 (0.551)

Tabela 9: Suscetibilidade à alteração dos diversos parâmetros comumente utilizados como instrumentos de avaliação em ensaios clínicos envolvendo pacientes com artrite reumatóide.

Parâmetro	Melhora clínica média / erro padrão (n = 19)
Número de artic. inflam.	6.20
Rigidez matinal	5.70
EPM-ROM	4.85
HAQ em português	4.72
Mactar	2.80
Força de preensão	2.26
Dor	1.68
VHS	0.68

V. DISCUSSÃO

Os parâmetros clínicos utilizados para avaliar doenças reumáticas têm sido estudados há mais de meio século. Nos últimos anos, entretanto, ao se avaliar a atividade das doenças, a simplicidade tem sido substituída pela complexidade e conseqüentemente muita controvérsia tem sido observada (53).

Há 53 anos um canadense, Dr. Douglas Taylor da Universidade McGill de Montreal, propôs a primeira classificação capaz de avaliar o grau de acometimento decorrente da artrite crônica (36). É interessante frisar que tal classificação precedeu em aproximadamente 10 anos a famosa e ainda utilizada classificação funcional de Otto Steinbrocker (17). Nos últimos anos, porém, houve uma grande diversificação dos instrumentos propostos para avaliar as doenças reumáticas, assim como tem havido grande preocupação na verificação e demonstração das propriedades de medida dos novos instrumentos ou parâmetros propostos.

Como anteriormente citado, os instrumentos que avaliam globalmente ou especificamente alguns aspectos do estado de saúde podem ter como finalidade a discriminação, a predição e a avaliação de modalidades terapêuticas.

Um instrumento com finalidade discriminativa é utilizado para distinguir indivíduos ou grupo de indivíduos com determinada característica, sendo que a observação de tal característica não pode ser feita através de uma padronização definida, ou seja, não há um "gold standard".

O teste de quociente de inteligência (QI), por exemplo, pode ser categorizado como um instrumento discriminativo, permitindo a distinção de crianças com maior ou menor capacidade de aprendizado (52).

Um instrumento com finalidade preditiva é utilizado com o intuito de identificar "a priori" quais indivíduos irão ou não desenvolver ou evoluir para determinadas características clínicas de uma doença. Estes instrumentos são freqüentemente utilizados na avaliação do prognóstico de diversas doenças. Como exemplo pode ser citado a medida de avaliação funcional que tem-se mostrado uma das medidas mais precisas na avaliação do prognóstico de pacientes com cancer (52).

Um instrumento com finalidade de avaliação é utilizado para medir a magnitude de sua alteração ao longo de um determinado tempo. Diversos instrumentos têm sido propostos para avaliar a variação da qualidade de vida ou capacidade funcional de pacientes reumáticos submetidos à diversas modalidades terapêuticas.

Ainda, assim como os objetivos destes 3 tipos de instrumentos são distintos, a demonstração de suas propriedades de medida também o são. Um instrumento válido para predição provavelmente não o é para avaliação ou discriminação, ou vice-versa. A citar, por exemplo, a propriedade suscetibilidade à alteração de um instrumento, altamente importante para um instrumento com finalidade de

avaliação e de verificação dispensável nos instrumentos com finalidade de predição e discriminação (52).

A seleção do instrumento de acordo com o objetivo de seu uso, bem como a verificação de suas propriedades de medida são de extrema importância na planificação de qualquer observação ou pesquisa clínica.

O HAQ, como definido pelos seus próprios autores, é um instrumento que foi desenvolvido basicamente para permitir a avaliação de modalidades terapêuticas propostas para pacientes com AR e outras doenças reumáticas. Na seleção de seus diversos componentes e questões algumas fontes foram pesquisadas, tais como o Arquivo Uniformizado das Doenças Reumáticas (54), a Escala de Convery e cols. (43) e as atividades da vida diária propostas por diferentes referências bibliográficas (39,42). As questões que avaliavam intuitivamente a capacidade funcional foram cuidadosamente estudadas e selecionadas.

Um questionário preliminar foi elaborado para ser administrado a pacientes por enfermeiras, através de uma entrevista. Quarenta pacientes com AR foram inicialmente entrevistados e questões imprecisas e ambíguas foram identificadas e modificadas. As questões foram então agrupadas em 8 componentes (vestir-se e cuidar-se, levantar-se, comer, caminhar, higiene, alcançar, preensão e outras atividades), sendo que cada um destes componentes era composto por 2 ou 3 questões. Uma escala tipo Lykert foi

utilizada na obtenção dos resultados individuais e do escore global do instrumento.

Na avaliação da reprodutibilidade do HAQ em inglês, Fries e cols (45) selecionaram e entrevistaram 20 pacientes com AR por 2 vezes num intervalo máximo de 12 dias. Duas versões do HAQ (auto-administrável e administrável através de uma entrevista) foram utilizadas. Em 10 pacientes a entrevista antecedeu a auto-administração e nos outros 10 a auto-administração antecedeu a entrevista. Um mesmo entrevistador administrou o questionário aos 20 pacientes. O teste estatístico utilizado na avaliação da reprodutibilidade foi o coeficiente de correlação de Spearman. As correlações observadas entre as 2 administrações para cada um dos componentes do instrumento variou de 0.56 a 0.85 (para o instrumento como um todo foi de 0.85), valores estes considerados altamente satisfatórios.

A validade do HAQ foi avaliada através da observação domiciliar da capacidade dos pacientes em executar as atividades da vida diária incluídas no questionário. Para tanto um avaliador observou diretamente as dificuldades enfrentadas pelos pacientes na execução destas atividades. O HAQ foi enviado pelo correio a 28 pacientes com AR com a solicitação que o mesmo fosse respondido adequadamente. Estes pacientes foram ainda avisados que poucos dias após o recebimento do questionário seriam visitados por um

avaliador para interrogá-lo sobre suas dificuldades na execução das tarefas da vida diária. Vinte e cinco pacientes concordaram com a visita domiciliar e foram avaliados. O intervalo entre a auto-administração do HAQ e a observação direta das dificuldades na execução das tarefas variou de 4 a 16 dias. Em 59% das respostas houve uma concordância perfeita entre o valor dos componentes individuais do HAQ relatado pelo paciente e o valor dado pela avaliação do observador. O teste estatístico Kappa ponderado, ao ser aplicado aos valores obtidos através das 2 avaliações, foi de 0.52, que pode ser interpretado como uma concordância moderada. O coeficiente de correlação de Spearman aplicado a cada componente variou de 0.47 a 0.88, valores estes estatisticamente significantes (45).

Ainda a acrescentar o fato de que após sua publicação o HAQ tem sido utilizado como um dos parâmetros clínicos de avaliação em ensaios terapêuticos envolvendo pacientes com AR, e conseqüentemente algumas de suas propriedades podem ser avaliadas indiretamente. Em 1986, Bombardier e cols. (55) ao executar um ensaio clínico para avaliar a efetividade do auranofin na AR, incluiu neste ensaio diversos parâmetros clínicos tradicionais, bem como alguns instrumentos para avaliar o estado de saúde e a qualidade de vida dos pacientes que foram estudados. Neste ensaio entrevistadores foram treinados para uma administração uniforme dos diversos questionários, entre estes o HAQ. Dos

311 pacientes com AR selecionados 294 completaram os 6 meses do estudo. As alterações observadas através do HAQ evidenciaram uma melhora estatisticamente significativa no grupo tratado com auranofin quando comparado com as alterações observadas no grupo tratado com placebo. Os outros parâmetros clínicos tradicionais (número de articulações dolorosas, número de articulações edemaciadas, tempo na caminhada de 15 metros e força de preensão) também constaram uma melhora estatisticamente significativa favorecendo o grupo tratado com auranofin. Este dado, embora simples, proporciona uma estimativa da validade construtiva longitudinal do HAQ. Uma análise mais cuidadosa dos dados já obtidos através deste ensaio clínico permitiria a definitiva constatação desta propriedade (56).

A suscetibilidade à alteração do HAQ em relação aos demais parâmetros clínicos utilizados em ensaios clínicos envolvendo pacientes com AR pode também ser avaliada através do ensaio clínico realizado por Bombardier e cols. (55). Assim, utilizando-se os dados publicados no referido estudo, a relativa suscetibilidade à alteração do HAQ (= 6.2) foi algo semelhante à relativa suscetibilidade à alteração dos parâmetros número de articulações edemaciadas (= 7.3) e número de articulações dolorosas (= 7.8), parâmetros estes, uma vez mais, considerados como melhores medidas de avaliação em ensaios envolvendo pacientes com AR (56).

A constatação das propriedades de medida do HAQ por

diversos autores (55,57-60), a verificação destas propriedades através dos dados acima apresentados, bem como sua administração simples e conveniente justificam a escolha deste instrumento como um parâmetro de avaliação clínica em ensaios envolvendo pacientes com AR.

Como anteriormente mencionado, para que traduções de um idioma para outro, de uma cultura para outra, sejam adequadamente realizadas, é necessária a utilização de termos que fazem referências a experiências reais familiares a ambas culturas. Se as características culturais são muito diferentes em termos de seus arranjos sociais e modos de viver, é extremamente difícil a obtenção de uma equivalência no significado de algumas expressões linguísticas, independentemente do quão cuidadoso a tradução é feita sob ponto de vista da linguagem. Problemas de equivalência ocorrem com maior frequência quanto menor é a similaridade cultural e linguística, portanto, a primeira fase do presente estudo objetivou a identificação e substituição das questões sem equivalência cultural.

Cinco questões foram identificadas: "Abrir uma caixa de leite" foi substituída por "Abrir um saco plástico de leite", em virtude das caixas de leite tipo longavida, embora disponíveis, são mais caras e conseqüentemente somente adquiridas por uma pequena proporção da população. "Tomar um banho de banheira" foi modificada para "Tomar um banho de chuveiro". As banheiras atualmente não são

freqüentemente encontradas em nossas residências. As questões "Abrir a porta de carros" e "Entrar e sair de um carro" foram substituídas respectivamente por "Segurar-se em pé no ônibus ou metrô" e "Entrar e sair de um ônibus", uma vez que a maioria dos pacientes brasileiros não possuem carros próprios e conseqüentemente só raramente podem executar tais atividades. "Executar tarefas como usar o aspirador de pó ou cuidar do jardim" foi modificada para "Executar tarefas como varrer ou limpar o chão". O brasileiro de modo geral raramente cuida do próprio jardim; normalmente solicita ajuda de um jardineiro quando quer ter um jardim bem cuidado, assim como há algumas famílias que não possuem o aspirador de pó.

É interessante notar que 2 das 5 questões (no. 7 e no. 20) que foram modificadas na versão em português do HAQ também foram modificadas na versão sueca do HAQ (61).

A reprodutibilidade intra-observador de cada um dos componentes da versão em português do HAQ, avaliada pelo coeficiente de correlação de Pearson variou de 0.501 a 0.793, enquanto a reprodutibilidade inter-observador destes componentes variou de 0.599 a 0.779. Os coeficientes de correlação do escore global da versão em português do HAQ foram 0.905 (intra-observador) e 0.830 (inter-observador), ambos valores considerados clinicamente satisfatórios. Como esperado, os coeficientes de correlação foram na maioria das vezes maiores na avaliação da reprodutibilidade intra-

observador (tabela 4). A reprodutibilidade intra-observador e inter-observador da versão em português do HAQ pode desta forma ser adequadamente constatada.

A coerência interna da versão em português do HAQ pode ser avaliada através da correlação entre as diversas questões que o compõem. Assim, os coeficientes de correlação de Pearson foram calculados (Tabela 5) com a finalidade de avaliar como os diferentes componentes presentes na versão em português se correlacionavam entre si. De forma semelhante a observada no HAQ em inglês, os coeficientes de correlação variaram de 0.285 a 0.774, com a maioria dos valores entre 0.400 e 0.600, correlação esta considerada moderada. Estes dados são importantes uma vez que algumas questões foram modificadas e a presença de uma alta correlação poderia sugerir a possibilidade de redundância entre os componentes. Por outro lado, a presença de um componente correlacionando-se muito discretamente ou não se correlacionando com outros componentes poderia sugerir que este tal componente estivesse avaliando preferentemente outra dimensão que não a de interesse, no caso a capacidade funcional. Este componente deveria desta forma também ser excluído do instrumento.

Exitem pelo menos 5 modos de se demonstrar a validade de um instrumento. A demonstração de algumas destas formas de validade são importantes para um tipo de instrumento ou índice e não para outros. Podemos citar: 1. Validade

aparente, que avalia se o instrumento realmente estuda a dimensão que se propõe; 2. Validade de conteúdo, que verifica se os componentes inclusos no instrumento são adequados para seu pretenso uso; 3. Validade do critério, que avalia a precisão do instrumento comparando o resultado obtido através do instrumento com outro parâmetro tido como padrão ou verdadeiro; 4. Validade discriminativa, que verifica se o instrumento é capaz de diferenciar estados de saúde tidos "a priori" como distintos; e 5. Validade construtiva, que verifica se o resultado obtido através do instrumento se relaciona com outros parâmetros de medida, uma vez que em algumas áreas da medicina, em especial na reumatologia, não há um parâmetro de medida único e perfeito capaz de avaliar o objeto de interesse (62).

O fato de o HAQ em inglês ter sido até então utilizado com frequência por investigadores com a finalidade de avaliar a melhora da capacidade funcional em pacientes com AR sugere que o mesmo tenha a validade aparente ("face validity"). Da mesma forma, em virtude do HAQ em inglês ter sido cuidadosamente construído utilizando-se diversas fontes bibliográficas e avaliado quanto a presença de questões ambíguas e redundantes, uma seleção adequada das questões e componentes foi obtida. Tal procedimento permite a observação de sua validade de conteúdo. Em decorrência das alterações executadas na versão em português do HAQ não serem estruturais é plausível pensar-se que a validade

aparente e de conteúdo foram mantidas mesmo após a tradução do HAQ. Estes 2 tipos de validade, no entanto, constituem-se em uma insuficiente evidência de validade ao se estudar um instrumento com finalidade de avaliação.

Recentemente, em um estudo de Wolfe e cols. (63), 400 pacientes com diagnóstico de AR, de acordo com os critérios da ARA de 1957, responderam ao HAQ. Os pacientes estudados foram agrupados de acordo com o resultado do HAQ em 3 grupos distintos, ou seja, escore do HAQ entre 0 e 1, 1.01 e 2, e 2.01 e 3. Os autores observaram que pacientes com resultados do HAQ mais elevados, isto é, com maior limitação funcional, apresentaram um maior número de articulações inflamadas ($P<0.001$), uma menor força de preensão ($P<0.001$), mais dor ($P<0.001$), e ainda eram mais deprimidos e ansiosos ($P<0.001$), quando comparados com os pacientes com uma menor limitação funcional. Observaram também que o grupo de pacientes com maior limitação funcional apresentava de forma geral um pior nível educacional.

Acrescenta-se ainda que a propriedade validade construtiva longitudinal foi satisfatoriamente observada tanto no HAQ em inglês, como na sua versão em português, ao se correlacionar a variação dos resultados do HAQ com a variação dos demais parâmetros clínicos comumente utilizados em ensaios clínicos envolvendo pacientes com AR.

A importância da demonstração da propriedade suscetibilidade à alteração nos instrumentos com finalidade

de avaliação, reside no fato destes instrumentos serem utilizados como o próprio nome o define como parâmetros de avaliação em ensaios clínicos; assim, a utilização de instrumentos altamente responsivos permite a redução da amostragem a ser estudada, ou seja, do número de pacientes a serem incluídos nos grupos de tratamento em investigação. Tal fato de certa forma não só viabiliza a realização destes ensaios clínicos, onde freqüentemente a disponibilidade e seleção de pacientes é restrita, como também contribui para tornar o estudo econômico.

Contrariamente às avaliações de reprodutibilidade e validade, não há até hoje um consenso quanto ao melhor modo estatístico de se avaliar a suscetibilidade à alteração de um determinado instrumento. Guyatt e cols. (64) sugerem que a medida da suscetibilidade à alteração seja feita através da razão da diferença mínima do escore do instrumento considerado como "cl clinicamente significativa" após uma intervenção terapêutica pela variabilidade dos resultados do instrumento em indivíduos clinicamente estáveis, ou seja, através de outra amostra de estudo independente da utilizada na avaliação da intervenção terapêutica. Esta razão pode dessa forma ser caracterizada como uma expressão da "magnitude do benefício", entretanto, por utilizar amostragens distintas em seu numerador e denominador, seu uso pode ser criticado em virtude da possibilidade de ocorrência de tendenciosidades (65).

Outros métodos propostos, tais como os baseados na teoria da generalização (66) e curva operacional ("ROC curves") (67) apresentam um valor controverso e conseqüentemente têm sido pouco utilizados.

Embora a suscetibilidade à alteração do HAQ em inglês não tenha sido até então sistematicamente avaliada, a utilização dos dados de ensaios clínicos que utilizaram o HAQ, como um dos parâmetros clínicos nos proporciona a possibilidade da determinação da relativa suscetibilidade à alteração deste instrumento. Desta forma tanto o HAQ em inglês, como a sua versão para o português apresentaram suscetibilidades à alteração semelhantes e satisfatórias.

Ultimamente, devido ao reconhecimento da importância da avaliação dos instrumentos para avaliar o estado de saúde ou a qualidade de vida, e conseqüentemente seu crescente uso em reumatologia, alguns autores têm proposto a utilização destes instrumentos não só na pesquisa clínica, como também na avaliação de rotina de pacientes reumáticos. Tal avaliação ambulatorial rotineira tem sido criticada, talvez, por ser complexa, demorada e muitas vezes pouco estimulantes tanto do ponto de vista dos pacientes como dos profissionais da área médica. Em contrapartida, Wolfe e cols. (63) chamam a atenção para o uso rotineiro do HAQ a nível ambulatorial, destacando a relevância clínica das informações obtidas através de sua administração. Estes autores afirmam que o resultado do HAQ é capaz de resumir e traduzir de forma

simples e única os valores observados pelos diversos parâmetros clínicos tradicionais, tais como, força de preensão, rigidez matinal, dor, número de articulações inflamadas, entre outros. Assim, em virtude da administração do HAQ ser fácil e rápida, bem como de baixo custo, o mesmo deveria ser introduzido como um parâmetro de avaliação clínica de rotina.

Estudos recentes têm também verificado e confirmado a acertiva de que a medida do estado de saúde se constitui em importante fator de predição de mortalidade em pacientes com AR (68-70). O HAQ poderia desta forma ser utilizado na predição da freqüência de utilização de serviços médicos e da mortalidade (63), dados estes de extremo interesse na planificação da conduta terapêutica de cada paciente.

A demonstração das propriedades de medida do HAQ como um instrumento de predição necessitam, entretanto, ser completamente avaliados e seus dados reproduzidos por diferentes autores.

Para finalizar é importante frisar que o objetivo principal de qualquer intervenção terapêutica na área médica consiste na obtenção e observação de uma melhora através dos diversos parâmetros clínicos e laboratoriais disponíveis na avaliação de uma determinada doença.

Atualmente dispomos de uma tecnologia capaz de avaliar a função imunológica, os constituintes do líquido articular, o estado anatômico da articulação, etc, porém há poucos

instrumentos capazes de adequadamente avaliar o impacto da doença na qualidade de vida de cada paciente. A força de preensão, uma radiografia ou o resultado de uma velocidade de hemossedimentação são incapazes de descrever a capacidade de um paciente em executar determinadas tarefas da vida diária.

O desenvolvimento de instrumentos ou parâmetros reprodutíveis, válidos e responsivos se constitui em uma área importante da pesquisa médica, pois a avaliação de qualquer procedimento médico depende do uso destes instrumentos ou parâmetros. A seleção e composição final de cada instrumento deve, portanto, estar diretamente relacionado ao objetivo do uso do proposto instrumento e às características da população a ser estudada.

Vale a pena ainda ressaltar, no entanto, que novos instrumentos não devem ser propostos até que uma cuidadosa revisão da literatura tenha sido realizada e os instrumentos localizados tenham sido criticamente avaliados. A demonstração das propriedades de medida de um determinado instrumento depende da avaliação e reavaliação através de seu uso em diferentes centros de pesquisa, por diferentes pesquisadores. Outro aspecto que deve ser ressaltado refere-se ao fato de que os resultados de diferentes centros de pesquisa só podem ser comparados quando tais pesquisas utilizam pacientes, intervenções terapêuticas e parâmetros clínicos semelhantes, senão idênticos.

A tradução para o português do instrumento HAQ para avaliar a capacidade funcional e sua adequação as condições sócio-econômicas e culturais do povo brasileiro, bem como a demonstração de suas propriedades de medida permitem a sua utilização como um valioso instrumento de avaliação em ensaios terapêuticos envolvendo pacientes brasileiros com AR.

VI. CONCLUSÕES

1. As propriedades de medida do HAQ (reprodutibilidade, validade e suscetibilidade à alteração) estão demonstradas o que justifica seu amplo uso na área da pesquisa reumatológica.
2. Ao serem traduzidas para o português 5 das 20 questões originariamente presentes no HAQ tiveram que ser modificadas com o intuito de se obter equivalência experimental.
3. A versão em português do HAQ é de fácil e rápida administração.
4. As propriedades de medida da versão em português do HAQ (reprodutibilidade, validade e suscetibilidade à alteração) foram demonstradas.
5. O HAQ em português se constitui em um valioso instrumento de avaliação em ensaios terapêuticos envolvendo pacientes brasileiros com AR.

VII. SUMMARY

Functional ability evaluation constitutes an important outcome measurement in any proposed trial involving patients with rheumatoid arthritis (RA). A crosscultural study was performed directed at the translation into Portuguese of the Health Assessment Questionnaire (HAQ) and the evaluation of its measurement properties. Five questions were modified in the Portuguese version of the HAQ to suit Brazilian conditions. The test-retest correlation coefficient was 0.905 ($p < 0.001$) and the inter-observer correlation coefficient was 0.830 ($p < 0.001$). The instrument kept its face, content and longitudinal construct validity. The responsiveness of the Portuguese version of the HAQ is somewhat similar to the one presented by the number of tender and swollen joints, measurement considered as the most preferable endpoint in RA trials. These data allow the use of the Portuguese version of the HAQ as a valuable evaluative instrument.

VIII. REFERÊNCIAS BIBLIOGRÁFICAS

1. Boyle JA, Buchanan WW: Clinical Rheumatology. Oxford, Blackwell Scientific Publications, 74, 1971.
2. Short CL: The Antiquity of Rheumatoid Arthritis. *Arthritis Rheum* 1974, 17:193.
3. Hartulung EF: Arthritis. The History and Conquest of Common Diseases, Edited by WR Bett, A Norman, University of Oklahoma Press, 132, 1954.
4. Snorrranson E: Landre-Beauvais and his Goutte Astenique Primitive. *Acta Med Scand* 1952, 142:Suppl 266:115-118.
5. Beighton P, Solomon L, Valkenburg H: Rheumatoid Arthritis in a rural South African negro population. *Ann Rheum Dis* 1975, 34:136-141.
6. Laine VAI: Rheumatic complaints in a urban population in Finland. *Acta Rheum Scand* 1962, 8:81-88.
7. Sato EI, Shichikawa K, Atra E, Inoue K, Takenaka Y: Estudo da prevalencia da artrite reumatoide em

populacao de origem japonesa em Mogi das Cruzes, Sao Paulo. Rev Bras Reum 1990, em impressao.

8. Seda H: Reumatismos, Conceito, Classificacao e Epidemiologia. In: Reumatologia Basica. Editado por Seda, H., 2^a Edicao, Capitulo 1, Editora Cultura Medica, 1982.
9. Medsger TAJr, Masi AT: Epidemiology of the Rheumatic Diseases. In: Arthritis and Allied Conditions. Edited by McCarty, D.J., tenth edition, Chapter 2, Lea & Febiger, 1985.
10. Adler E, Abranson JH, Elkan Z, Ben Hador S, Goldenberg R: Rheumatoid Arthritis in a Jerusalem population 1. Epidemiology of the disease. Am J Epidemiol 1967, 85:365-377.
11. Mendez-Bryan R, Gonzales-Alcover R, Roger L: Rheumatoid Arthritis: Prevalence in a tropical area. Arthritis Rheum 1964, 7:171-176.
12. Williams RCJr, McCarty DJ: Clinical Picture of Rheumatoid Arthritis. In Arthritis and Allied Conditions. Edited by McCarty, D.J., tenth edition, Chapter 38, Lea & Febiger, 1985.

13. Silver RM, Zvaifler NJ: Immunopathogenesis of Rheumatoid Arthritis. Edited by P.D. Utsinger, N.J. Zvaifler and G.E. Erlich, Philadelphia, J.B. Lippincott, 1984.
14. Startny P: Association of the B-cell alloantigen DR_{w4} with Rheumatoid Arthritis. *N Engl J Med* 1978, 298:896.
15. Ritchie DM, Boyle JA, McInnes JM et al.: Clinical studies with an articular index for the assessment of joint tenderness in atients with rheumatoid arthritis. *Q J Med* 1968, New Series XXXVII 147:393-406.
16. Lansbury J: Report of a three-year study on the systemic and articular indexes in rheumatoid arthritis - theoretic and clinical considerations. *Arthritis Rheum* 1958, 1:505-522.
17. Steinbrocker O, Traeger CH, Batterman RC: Therapeutic criteria in rheumatoid arthritis. *J Am Med Assoc* 1949, 140(8):659-662.
18. Sherrer YS, Block DA, Mitchell DM, Young DY, Fries JF: The developmente of disability in rheumatoid

arthritis. *Arthritis Rheum* 1986, 29:494-500.

19. Colton R: The natural course of rheumatoid arthritis. *Minn Med* 1979, 62:129-130.
20. Meenan RF, Yelin EH, Henke DJ, Curtis DL, Epstein WV: The cost of rheumatoid arthritis: a patient-oriented study of chronic disease costs. *Arthritis Rheum* 1978, 21:827-833.
21. Liang MH, Larson M, Thompson M: Costs and outcomes in rheumatoid arthritis and osteoarthritis. *Arthritis Rheum* 1984, 27:522-529.
22. Lubeck DP, Sitz PW, Fries JF, Wolfe F, Mitchell D, Roth SH: A multi-center study of utilization and costs in rheumatoid arthritis. *Arthritis Rheum* 1986, 29:488-493.
23. Yelin EH, Henke CJ, Kramer JS, Nevitt M, Epstein WV: A comparison of the treatment of rheumatoid arthritis in health maintenance organizations and fee-for-service practices. *N Engl J Med* 1985, 15:962-967.

24. World Health Organization. International classification of impairments, disabilities and handicaps. A manual classification relating to the consequences of disease. Geneva. World Health Organization, 1980.
25. Bergner M, Bobbitt RA, Carter WB et al.: The Sickness Impact Profile: development and final revision of a health status measure. *Med Care* 1981, 19:787-805.
26. Ott CR, Sivarajan ES, Newton KM et al.: A controlled randomized study of early cardiac rehabilitation: the Sickness Impact Profile as an assessment tool. *Heart Lung* 1983, 12:162-170.
27. Liang MH, Larson MG, Cullen KE et al.: Comparative measurement efficiency and sensitivity of five health status instruments for arthritis. *Arthritis Rheum* 1985, 28:542-547.
28. Deyo RA, Diehl AK, Rosenthal M: How many days of bed rest for acute low back pain? A randomized clinical trial. *N Engl J Med* 1986, 315:1064-1070.

29. Hunt SM, McKenna SP, McEwen J et al.: A quantitative approach to perceived health status: a validation study. *J Epidemiol Community Health* 1980, 34:281-286.
30. Chambers LW, Mc Donald LA, Tugwell P et al.: The McMaster Health Index Questionnaire as a measure of quality of life for patients with rheumatoid arthritis. *J Rheumatol* 1982, 9:780-784.
31. Torrance GW: Utility approach to measuring health-related quality of life. *J Chron Dis* 1987, 40: 593-600.
32. Sackett DL, Torrance GW: The utility of different health states as perceived by the general public. *J Chron Dis* 1978, 31:697-704.
33. Guyatt GH, Townsend M, Pugsley SO et al.: Bronchodilators in chronic airflow limitation: effects on airway function, exercise capacity and quality of life. *Am Rev Respir Dis* 1987, 35:1069-1074.
34. Guyatt GH, Berman LB, Townsend M et al.: A measure of quality of life for clinical trials in chronic

lung disease. *Thorax* 1987, 42:773-778.

35. Olsson G, Lubsen J, van Es G et al.: Quality of life after myocardial infarction: effect of long term metoprolol on mortality and morbidity. *Br Med J* 1986, 292:1491-1493.
36. Taylor D: A table for the degree of involvement in chronic arthritis. *Can Med Assoc J* 1937, 36:608-610.
37. Empire Rheumatism Council. Multi-centre controlled trial comparing cortisone acetate and acetyl salicylic acid in the long-term treatment of rheumatoid arthritis. *Ann Rheum Dis* 1955, 14:353-370.
38. Donaldson SW, Wagner CC, Gresham GE: A unified ADL evaluation form. *Arch Phys Med Rehabil* 1973, 54:175-179.
39. Katz S, Downs TD, Cash HR, Grotz RC: Progress in development of the index of ADL. *Gerontologist* 1970, 10:20-30.

40. Lee P, Jasami MK, Dick WC, Buchanan WW: Evaluation of functional index in rheumatoid arthritis. *Scand J Rheumatol* 1973, 2:71-77.
41. Ebert DR, Fashing V, Rahifs V, Schteyer J, Wolf R: Repeatability and objectivity of various measurements in rheumatoid arthritis: a comparative study. *Arthritis Rheum* 1976, 19:1278-1286.
42. Mahoney FI, Barthel DW: Functional evaluation: the Barthel Index. *Md State Med J* 1965, 14:61-65.
43. Convery FR, Minter MA, Amiel D, Conneti KL: Polyarticular disability: a functional assessment. *Arch Phys Med Rehabil* 1977, 58:494-499.
44. Jette AM: Functional Status Index: reliability of a chronic disease evaluation instrument. *Arch Phys Med Rehabil* 1980, 61:395-401.
45. Fries JF, Spitz P, Kraines RG, Holman AR: Measurement of patient outcome in arthritis. *Arthritis Rheum* 1980, 23:137-145.

46. Kaplan RM, Bush JW, Berry CC: Health Status: types of validity for an index of well-being. *Health Serv Res* 1976, 11:478-507.
47. Meenan RF, Gertman PM, Mason JM: Measuring health status in arthritis: The Arthritis Impact Measurement Scale. *Arthritis Rheum* 1980, 23:146-152.
48. Tugwell P, Bombardier C, Buchanan WW, Goldsmith CH, Grace E, Hanna B: The Mactar patient preference questionnaire - An individual functional priority approach for assessing improvement in physical disability in clinical trial in rheumatoid arthritis. *J Rheumatol* 1987, 14: 446-457.
49. Ferraz MB, Oliveira LM, Araujo PM, Atra E: EPM-ROM scale: an evaluative instrument to be used in rhaumetoid arthritis. *Clin Exp Rheumatol* 1990, 8:491-494.
50. Secherest L, Fay TL, Zaidi SMH: Problems of translation in cross-cultural research. *J Cross-Cultural Psychol* 1972, 3:41-56.

51. Werner O, Campbell DT: Translating, working through interpreters and the problem of decentering. In A handbook of methods in cultural anthropology. Edited by Navon, G & Cohen J. The Natural History Press, 398-420. New York, 1970.
52. Kirshner B, Guyatt G: A methodological framework for assessing health indices. *J Chron Dis* 1985, 38:27-36.
53. Bellamy N: Critical review of clinical assessment techniques for rheumatoid arthritis trials: new developments. *Scand J Rheumatol* 1989, (Suppl 80):3-16.
54. Fries JF, Hess EV, Klinenberg J: A standard data base for rheumatic disease. *Arthrtis Rheum* 1974, 17:327-336.
55. Bombardier C, Ware J, Russell D, et al.: Auranofin therapy and quality of life in patients with rheumatoid arthritis, results of a multicenter trial. *Am J Med* 1986, 81:565-578.
56. Ferraz MB, Atra E: Rheumatoid arthritis and the measurement properties of the physical ability

dimension of the Stanford Health Assessment Questionnaire. Clin Exp Rheumatol 1989, 7:341-344.

57. Liang MH, Larson MG, Cullen KE, Schwartz JA: Comparative measurement, efficiency and sensitivity of five health status instruments for arthritis research. Arthritis Rheum 1985, 28:542-547.
58. Fries JF, Spitz PW, Young DY: The dimensions of health outcomes: The health assessment questionnaire, disability and pain scales. J Rheumatol 1982, 9:789-793.
59. Liang MH, Cullen K, Larson M: Search of a more perfect Mousetrap (Health status or Quality of life instrument). J Rheumatol 1982, 9:775-779.
60. Brown JH, Katis LE, Spitz PW, Gertman P, Fries JF, Meenan RF: The dimensions of health outcomes: a cross-validated examination of health status measurement. Am J Public Health 1984, 74:159-161.
61. Ekdahl C, Eberhardt K, Andersson SI, Svensson B: Assessing disability in patients with rheumatoid arthritis - Use of a Swedish version of the

Stanford Health Assessment Questionnaire. *Scand J Rheumatol* 1988, 17:263-271.

62. Bombardier C, Tugwell P: Methodological considerations in functional assessment. *J Rheumatol* 1978, (Suppl 15) 14:6-10.
63. Wolfe F, Kleinheksel SM, Cathey MA, Hawley DJ, Spitz PW, Fries JF: The clinical value of the SHAQ Functional disability index in patients with rheumatoid arthritis. *J Rheumatol* 1988, 15:1480-1488.
64. Guyatt G, Walter SW, Norman GR: Measuring change over time: Assessing the usefulness of evaluative instruments. *J Chron Dis* 1987, 40:171-178.
65. Norman GR: Issues in the use of change scores in randomized trials. *J Clin Epidemiol* 1989, 42:1097-1105.
66. Cronbach LJ, Gleser G, Nanda H, Rajaratnan N: The dependability of behavioral measurements for scores and profiles. New York. John Wiley; 1972.

67. Deyo RA, Centor RM: Assessing the responsiveness of functional scales to clinical change: an analogy to diagnostic test performance. *J Chron Dis* 1986, 39:897-906.
68. Kasis LE, Anderson JJ, Meenan RF: Health Status as a predictor of mortality in rheumatoid arthritis: a five year study. *J Rheumatol* 1990, 17:609-613.
69. Pincus T, Callahan LF, Vaughn WK: Questionnaire, walking-time and button test measures of functional capacity as predictive markers for mortality in rheumatoid arthritis. *J Rheumatol* 1987, 14:240-251.
70. Mitchell DM, Spitz PW, Young DY, Bloch DA, McShane DJ, Fries JF: Survival, prognosis and causes of death in rheumatoid arthritis. *Arthritis Rheum* 1986, 29:706-714.